Transfer Learning with BERT: Tweet sentiment on COVID-19 vaccine

Abstract

The COVID-19 pandemic has changed our lives forever. People are on social media to share their feelings and opinions during the pandemic. We can learn about public attitudes toward COVID-19 vaccines by looking at what's said on social media. However, there are not many studies.

In my research, I used several public datasets on Kaggle, including tweets about COVID-19 vaccination extracted from the Twitter API during the pandemic, along with a dataset with Twitter texts and corresponding sentiment labels. I conducted the sentiment analysis model through transfer learning, based on a Bert model published by Google and fine tuned with the labeled Twitter data.

Using my model, I looked at perceptions of different vaccine brands and how those perceptions have changed over time. By applying the sentiment analysis model, I saw that people's emotions towards vaccines gradually changed from the extreme at the beginning to the neutral. I also found that people in most countries were positive about most brands of COVID-19 vaccines, with BioNTech being the most positive.

Introduction

COVID-19 has quickly spread around the world, killing many lives. Pharmaceutical companies and universities have been working on COVID-19 vaccines to prevent further spread of the virus. Some vaccines are approved for the market. While people may have doutes at the beginning, vaccination is now in full swing in many countries. In order to effectively promote vaccines, it is important to collect people's opinions on vaccines. Understanding people's attitudes and willingness to get vaccinated could also help predict the spread of the pandemic.

Background

With advances in Machine Learning (ML), neural network-based methods, such as Convolutional/Recurrent Neural Networks, have been proposed to solve many NLP tasks [1]. Among many techniques, Bidirectional Encoder Representations from Transformers (BERT) [2], a type of pre-trained language models, and can generate dynamic word embeddings, shows state-of-the-art results on a wide array of NLP tasks, especially in sentiment analysis tasks of social media text.

BERT model usually contains a huge number of parameters that can range from 100 million to over 300 million. And training a BERT model from scratch on a small dataset would result in overfitting. pre-trained BERT model that was trained on a huge dataset, as a starting point. We can then conduct model fine-tuning by further training the model on our domain-specific dataset, which is usually relatively smaller in size.

Recent researches have shown that fine-tuning the pre-trained language model with limited annotated domain-specific data has achieved excellent performance in a series of NLP tasks. For example, Adhikari et al. [3] established state-of-the-art results for four accessible datasets (Reuters, AAPD, IMDB, Yelp 2014) by fine-tuning BERT for document classification.

Another example is, in the work of Azzouza et al. [4], a framework for twitter sentiment analysis based on BERT has been proposed and achieved good results. The framework achieved high performance on the SemEval 2017 dataset.

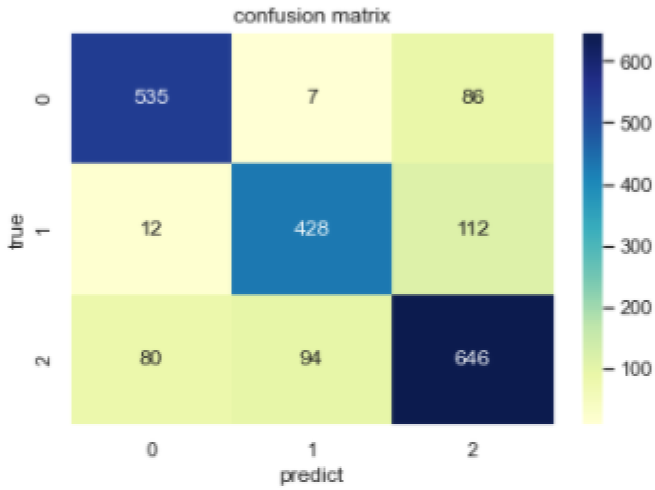
Methods

My sentiment analysis model consists of two parts.

In particular, I first used pre-labeled Twitter data to fine-tune the BERT model published by Google. The Model classifies the sentiment of Twitter posts into positive, neutral and negative categories.

After sentiment classification, the mean of the sentiment is calculated. And I looked into the trend of people’s average sentiment evolving over time, by brand of Vaccine and also by region.

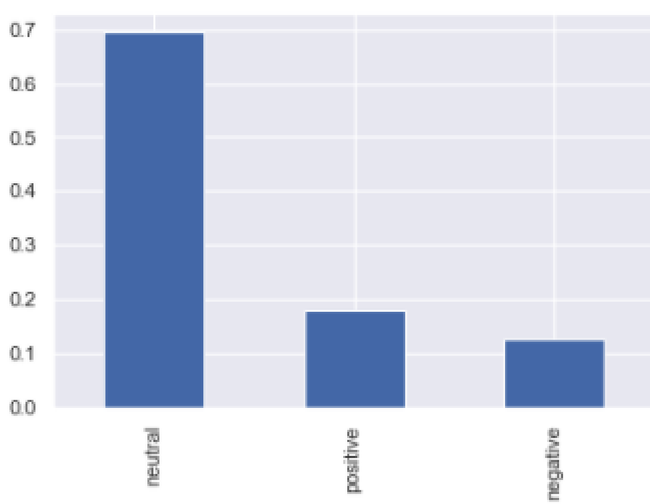
Results and discussion



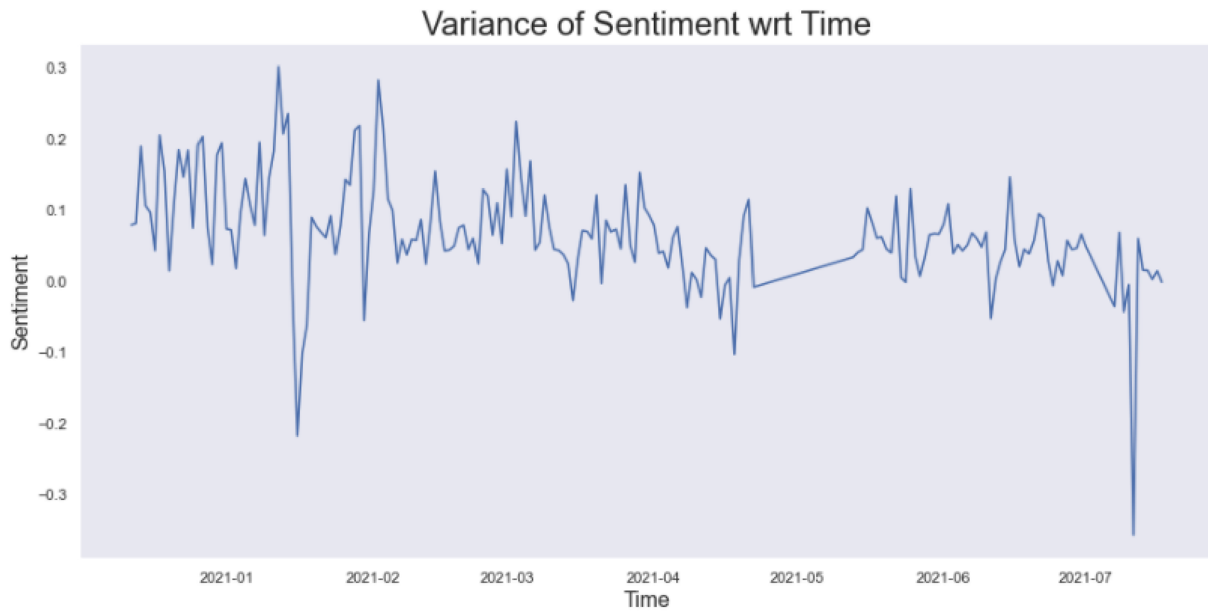
0.8045

	precision	recall	f1-score	support
0	0.8533	0.8519	0.8526	628
1	0.8091	0.7754	0.7919	552
2	0.7654	0.7878	0.7764	820
accuracy			0.8045	2000
macro avg	0.8092	0.8050	0.8070	2000
weighted avg	0.8050	0.8045	0.8046	2000

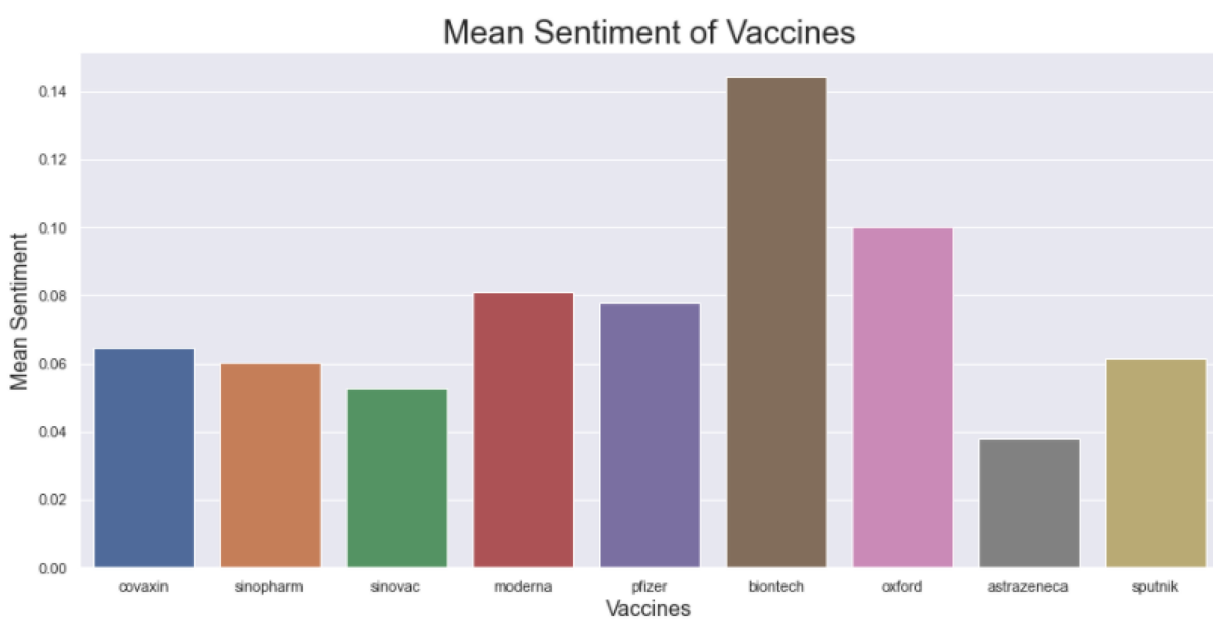
Bert model, fine tuned with Twitter data, is showing an accuracy of 0.8045(f1-score).



Around 70% of the Tweets on COVID-19 vaccination are neutral. About 18% are classified as positive and about 12% are believed to be leaning towards negative. This in general inline with the general distribution of peoples sentiment and emotions on things. Slightly more tweets are classified as positive shows that the emotions of the people on Twitter towards COVID-19 vaccination is positive.



As a common trend with all vaccines, we can see that the emotions at the initial stage seem to be bouncing between two extremes. The positive extreme can be attributed to the celebration of the discovery of the vaccine, while the negative extreme could be the result of widespread rumors about the vaccine's credibility. All vaccines that started with negative emotions also reached the extreme of positive emotions when rumours about them were quashed by concrete evidence from the authorities. As the vaccinations progressed, the mood turned more neutral.



The general emotions towards different brands of vaccines are all slightly positive, with BioNTech being the most positive.

Conclusion

My goal is to find an effective approach to perform sentiment analysis on Tweets on COVID-19 vaccines, with satisfying performance. I built a model that is enhanced by a transfer learning system, with pre-labeled Tweets data. It was conducted in two steps, where the first step involved a series of pre-processing procedures, and the second step exploited a version of BERT, which was pre-trained and published by Google and then fine-tuned in this research with Tweets text. In this way, I was able to take advantage of the strong feature extraction capability of large neural networks and also fine-tune the language model to fit the specific domain I'm researching in.

Future work will be directed to improve the pre-processing procedure to better clean up the Tweets text, such as emoji and retweet prefix. Another area to improve is that the data for fine tuning could be more domain specific. For example, we could use medication specific sentiment data.

Reference

1. Prima Dewi Purnamasari, Muhammad Taqiyuddin, Anak Agung Putri Ratna. *Performance comparison of text-based sentiment analysis using recurrent neural network and convolutional neural network.* ICCIP '17: Proceedings of the 3rd International Conference on Communication and Information ProcessingNovember 2017 Pages 19–23 [CrossRef]
2. Devlin J., Chang M., Lee K., Toutanova K. BERT: *Pre-training of Deep Bidirectional Transformers for Language Understanding*. In: Burstein J., Doran C., Solorio T., editors. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019; Minneapolis, MN, USA. 2–7 June 2019; Stroudsburg, PA, USA: Association for Computational Linguistics; 2019. pp. 4171–4186. [CrossRef] [Google Scholar]
3. Adhikari A., Ram A., Tang R., Lin J. DocBERT: BERT for Document Classification 2019. *arXiv.* 20191904.08398 [Google Scholar]
4. Azzouza N., Akli-Astouati K., Ibrahim R. *Advances in Intelligent Systems and Computing.* Volume 1073. Springer Science and Business Media LLC; Berlin/Heidelberg, Germany: 2019. TwitterBERT: Framework for Twitter Sentiment Analysis Based on Pre-trained Language Model Representations; pp. 428–437. [Google Scholar]